PATH VI: a pathsearch method for variational inequalities

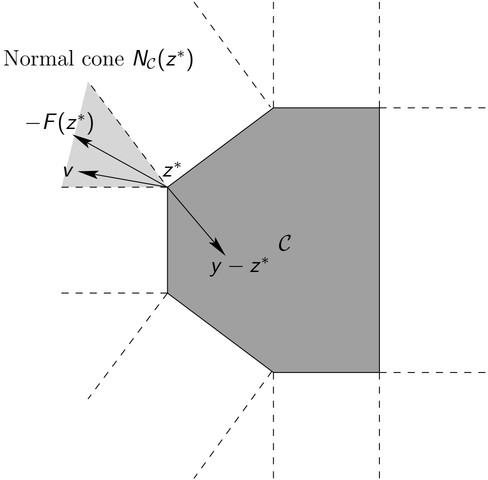
Michael C. Ferris
(Joint with Youngdae Kim and Todd Munson)

University of Wisconsin, Madison

Funded by DOE-MACS Grant with Argonne National Laboratory

Computational Contact Mechanics: Advances and Frontiers in Modeling Contact February 17, 2014

VI: $-F(z) \in \mathcal{N}_{\mathcal{C}}(z)$



Many applications where F is not the derivative of some f

Variational Inequality Formulation

- $F : \mathbb{R}^n \to \mathbb{R}^n$
- Ideally: $\mathcal{C} \subseteq \mathbb{R}^n$ – constraint set
- Often: $\mathcal{C} \subseteq \mathbb{R}^n$ – simple bounds

$$0 \in F(z) + N_{\mathcal{C}}(z)$$

- VI generalizes many optimization problems: LP, MCP, and LCP
  - For Nonlinear Equations: $F(z) = 0$ set $\mathcal{C} \equiv \mathbb{R}^n$
  - For NCP: $0 \leq F(z)$, $z \geq 0$ and $z^T F(z) = 0$ set $\mathcal{C} \equiv \mathbb{R}^n_+$
  - For LCP, set $F(z) = Mz + q$ and $\mathcal{C} \equiv \mathbf{R}^n_+$.
  - For MCP (rectangular VI), set $\mathcal{C} \equiv [l, u]^n$.
  - Example: convex optimization first-order optimality condition:

$$\min_{z \in \mathcal{C}} f(z) \iff -\nabla f(z) \in N_{\mathcal{C}}(z) \quad \iff \quad 0 \in \nabla f(z) + N_{\mathcal{C}}(z)$$

  - For LP, set $F(z) \equiv \nabla f(z) = p$ and $\mathcal{C} = \{z \mid Az = a, Hz \leq h\}$.

AVI over polyhedral convex set

An affine function

$$F : \mathbb{R}^n \to \mathbb{R}^n,\; F(z) = Mz + q,\; M \in \mathbb{R}^{n\times n},\; q \in \mathbb{R}^n$$

A polyhedral convex set

$$\mathcal{C} = \{z \in \mathbb{R}^n \mid Az(\geq, =, \leq)a,\; l \leq z \leq u\},\; A \in \mathbb{R}^{m\times n}$$

Find a point $z^* \in \mathcal{C}$ satisfying

$$\begin{aligned} \langle F(z^*), y - z^* \rangle \quad & \geq 0, \quad \forall y \in \mathcal{C} \\ (\Leftrightarrow)\ \langle -F(z^*), y - z^* \rangle \quad & \leq 0, \quad \forall y \in \mathcal{C} \\ (\Leftrightarrow)\ -F(z^*) \quad & \in N_{\mathcal{C}}(z^*) \end{aligned}$$

where

$$N_{\mathcal{C}}(z^*) \quad = \{v \mid \langle v, y - z^* \rangle \leq 0, \forall y \in \mathcal{C}\}$$

Variational inequalities (current state)

- Find $z \in \mathcal{C}$ such that

$$0 \in F(z) + \mathcal{N}_{\mathcal{C}}(z)$$

- model vi / F, g /;
  empinfo: vi F z g
- Convert problem into complementarity problem by introducing multipliers on representation of e.g. $\mathcal{C} = \{z \in [l,u] : g(z) \leq 0\}$

$$\begin{bmatrix} F(z) - \nabla g(z)\lambda \\ g(z) \end{bmatrix} + \mathcal{N}_{[l,u]\times\mathbb{R}^m_+}$$

- $\mathcal{C}$ polyhedral (e.g. $\mathcal{C} = \{z \in [l,u] : Az \leq a\}$ and $F(z) = Mz + q$

$$\begin{bmatrix} M & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} z \\ \lambda \end{bmatrix} + \begin{bmatrix} q \\ -a \end{bmatrix} + \mathcal{N}_{[l,u]\times\mathbb{R}^m_+}$$

Theorem

Suppose $\mathcal{C}$ is a polyhedral convex set and M is an L–matrix with respect to $\mathrm{rec}\mathcal{C}$ which is invertible on the lineality space of $\mathcal{C}$. Then exactly one of the following occurs:

- *PATHAVI solves (AVI)*
- *the following system has no solution*

$$Mz + q \in (\mathrm{rec}\mathcal{C})^D, \qquad z \in \mathcal{C}. \tag{1}$$

Corollary

If M is copositive–plus with respect to $\mathrm{rec}\mathcal{C}$, then exactly one of the following occurs:

- *PATHAVI solves (AVI)*
- *(1) has no solution*

Note also that if $\mathcal{C}$ is compact, then any matrix M is an L–matrix with respect to $\mathrm{rec}\mathcal{C}$. So always solved.

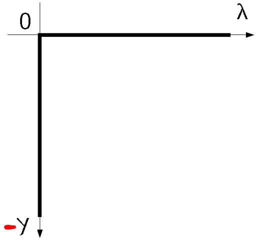
Experimental results: AVI vs MCP

PATH is a solver for MCP (mixed complementarity problem).

- Run PathAVI over AVI formulation.
- Run PATH over AVI in MCP form (poorer theory as $\mathrm{rec}\mathcal{C}$ larger).
- Data generation
  - M is an $n \times n$ symmetric positive definite/indefinite matrix.
  - A has m randomly generated bounded inequality constraints.

(m, n)	PathAVI		PATH		% negative eigenvalues
	status	# iterations	status	# iterations	
(180,60)	S	55	S	72	0
(180,60)	S	45	S	306	20
(180,60)	S	2	F	9616	60
(180,60)	S	1	F	10981	80
(360,120)	S	124	S	267	0
(360,120)	S	55	S	1095	20
(360,120)	S	2	F	10020	60
(360,120)	S	1	F	7988	80

Complementarity Problems via Graphs



$$\mathcal{T} = \mathcal{N}_{\mathbb{R}_+} = (\mathbb{R}_+ \times \{0\}) \bigcup (\{0\} \times \mathbb{R}_-)$$

$$-y \in \mathcal{T}(\lambda) \iff (\lambda, -y) \in \mathcal{T} \iff 0 \le \lambda \perp y \ge 0$$

By approximating (smoothing) graph can generate interior point algorithms for example $y\lambda = \epsilon, y, \lambda > 0$

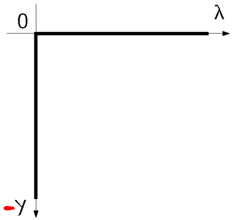
$$-F(z) \in \mathcal{N}_{\mathbb{R}^n_+}(z) \iff (z, -F(z)) \in \mathcal{T}^n \iff 0 \le z \perp F(z) \ge 0$$

Complementarity Systems (DVI)

$$\frac{dx}{dt}(t) = f(x(t), \lambda(t))$$

$$y(t) = h(x(t), \lambda(t))$$

$$0 \leq y(t) \perp \lambda(t) \geq 0$$

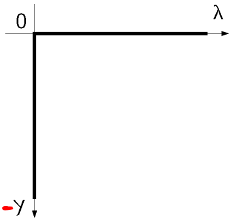


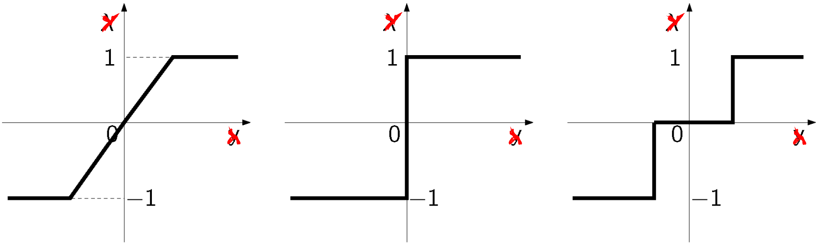
Complementarity Systems (DVI)

$$\frac{dx}{dt}(t) = f(x(t), \lambda(t))$$

$$y(t) = h(x(t), \lambda(t))$$

$$0 \leq y(t) \perp \lambda(t) \geq 0$$





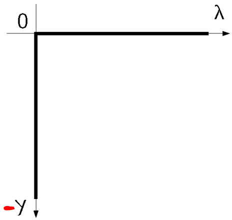
saturation Relay Relay with dead zone

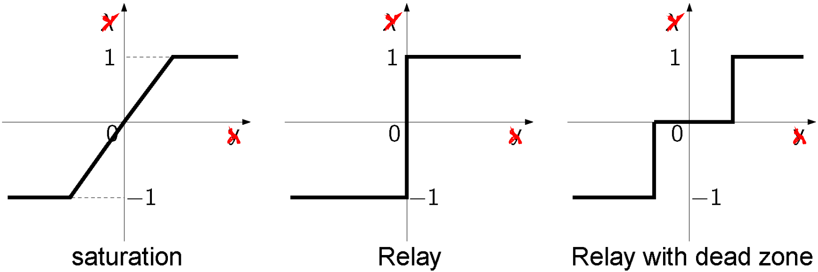
Complementarity Systems (DVI)

$$\frac{dx}{dt}(t) = f(x(t), \lambda(t))$$

$$y(t) = h(x(t), \lambda(t))$$

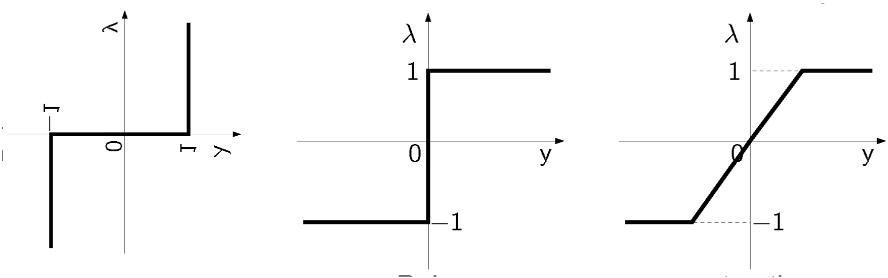
$$(\lambda(t), -y(t)) \in \mathcal{T}$$





Operators and Graphs ($\mathcal{C} = [-1, 1], \mathcal{T} = \mathcal{N}_{\mathcal{C}}$)

$$z_i = -1, -F_i(z) \leq 0 \text{ or } z_i \in (-1, 1), -F_i(z) = 0 \text{ or } z_i = 1, -F_i(z) \geq 0$$



$\mathcal{T}(\lambda)$ $\qquad$ $\mathcal{T}^{-1}(y)$ $\qquad$ $(\mathcal{I} + \mathcal{T})^{-1}(y) = P_{\mathcal{T}}(y)$

$P_{\mathcal{T}}(y)$ is the projection of y onto $[-1, 1]$

Generalized Equations

- Suppose $\mathcal{T}$ is a maximal monotone operator

$$0 \in F(z) + \mathcal{T}(z) \qquad (GE)$$

- Define $P_{\mathcal{T}} = (\mathcal{I} + \mathcal{T})^{-1}$
- If $\mathcal{T}$ is polyhedral (graph of $\mathcal{T}$ is a finite union of convex polyhedral sets) then $P_{\mathcal{T}}$ is piecewise affine (continous, single-valued, non-expansive)

$$\begin{aligned} 0 \in F(z) + \mathcal{T}(z) \quad &\iff \quad z \in F(z) + \mathcal{I}(z) + \mathcal{T}(z) \\ &\iff \quad z - F(z) \in (\mathcal{I} + \mathcal{T})(z) \iff P_{\mathcal{T}}(z - F(z)) = z \end{aligned}$$

Use in fixed point iterations (cf projected gradient methods)

Normal Map

- Suppose $\mathcal{T}$ is a maximal monotone operator

$$0 \in F(z) + \mathcal{T}(z) \quad (GE)$$

- Define $P_{\mathcal{T}} = (I + \mathcal{T})^{-1}$

$$\begin{aligned}
0 \in F(z) + \mathcal{T}(z) &\iff z \in F(z) + \mathcal{I}(z) + \mathcal{T}(z) \\
&\iff z - F(z) = x \text{ and } x \in (\mathcal{I} + \mathcal{T})(z) \\
&\iff z - F(z) = x \text{ and } P_{\mathcal{T}}(x) = z \\
&\iff P_{\mathcal{T}}(x) - F(P_{\mathcal{T}}(x)) = x \\
&\iff 0 = F(P_{\mathcal{T}}(x)) + x - P_{\mathcal{T}}(x)
\end{aligned}$$

This is the so-called Normal Map Equation

Key idea of algorithm $\mathcal{T} = \mathcal{N}_{\mathcal{C}}$

Homotopy: Easy solution for μ large, drive $\mu \to 0$.

$$\mu r = F(\pi_{\mathcal{C}}(x(\mu))) + x(\mu) - \pi_{\mathcal{C}}(x(\mu))$$

Define $z(\mu) = \pi_{\mathcal{C}}(x(\mu))$, then

$$\mu r = F(z(\mu)) + x(\mu) - z(\mu)$$

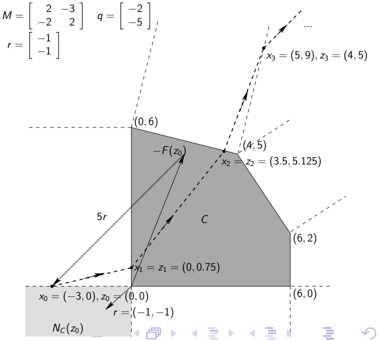
$$\begin{aligned}
x - z \quad & \in N_{\mathcal{C}}(z) \\
N_{\mathcal{C}}(z) \quad & = \{-A^T u - w + v\} \\
\text{such that} \quad & Az(\geq, =, \leq)a \;\perp\; u(\geq, \text{free}, \leq)0 \\
& 0 \leq w \;\perp\; z - l \geq 0 \\
& 0 \leq v \;\perp\; u - z \geq 0
\end{aligned}$$

Ray start and complementary pivoting

Solve the normal map by

1. Computing an extreme point $z_e \in \mathcal{C}$ by solving Phase I.
2. Introducing a ray with a covering vector r in the interior of the normal cone at z_e.
3. Setting up an initial basis for complementary pivoting using the result of Phase I.
4. Doing complementary pivoting until the multiplier on r becomes zero.

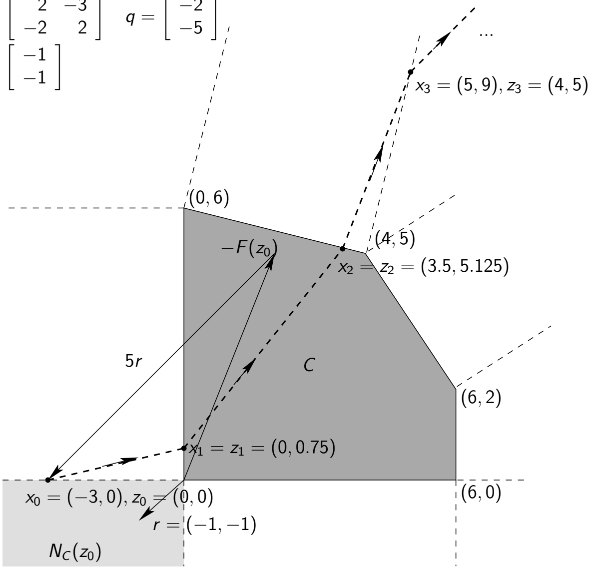
$$
\begin{aligned}
-(Mz+q)+\mu r &= -A^T u - w + v \\
Az(\geq,=,\leq)a &\perp u(\geq,\text{free},\leq)0 \\
0 \leq w &\perp z - l \geq 0 \\
0 \leq v &\perp u - z \geq 0 \\
\mu &\geq 0
\end{aligned}
$$



Example (complementary pivoting)

$$M = \begin{bmatrix} 2 & -3 \\ -2 & 2 \end{bmatrix} \quad q = \begin{bmatrix} -2 \\ -5 \end{bmatrix}$$

$$r = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$



Implementation

1. Solve Phase I over $\mathcal{C}$ using CPLEX.

$$\begin{aligned} \underset{z}{\text{minimize}} \quad & 0^T z \\ \text{subject to} \quad & Az = a \\ & l \le z \le u \end{aligned}$$

   - We have included slack and artificial variables.
   - Thus, rank $A = m$.

2. Do complementary pivoting (Lemke's method) until a feasible solution or a secondary ray is found.

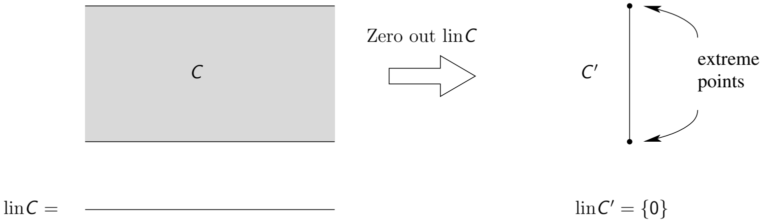
Large scale implementation: Computing an extreme point

No extreme point exists when C has a non-zero lineality space

$$\operatorname{lin} C = \ker \begin{bmatrix} A \\ H \end{bmatrix} \neq \{0\}$$

(H encodes bounds.) In that case, we compute a boundary point of C.

- Computing a boundary point of C
  - Zero out $\operatorname{lin} C$ and compute an extreme point over reduced space.



Solving Phase I

If feasible region of $\mathcal{C}$ is not empty, then CPLEX comes with a basis triple (B, N_l, N_u) with $\mathbf{B} = A_B$ nonsingular such that

- $B = (B_1, \ldots, B_m) \subseteq \{1, \ldots, n\}$: indices of basic variables
- $N = \{1, \ldots, n\} \setminus B$: indices of nonbasic variables
- $N_l \cap N_u = \emptyset, N_l \cup N_u = \{j \notin B : x_j$ neither fixed nor free$\}$, $l_j > -\infty$ for $j \in N_l$ and $u_j < +\infty$ for $j \in N_u$
- $N_{fr} = \{j \in N : z_j$ free$\}$ and $N_{fx} = \{j \in N : z_j$ fixed$\}$.
- Note that $z_{N_l} = l_{N_l}, z_{N_u} = u_{N_u}, z_{N_{fr}} = 0, z_{N_{fx}} = l_{N_{fx}} = u_{N_{fx}}$, and $z_B = \mathbf{B}^{-1}(b - A_N z_N)$.

Phase I result interpretation (when $\exists$ an extreme point)

If $N_{fr} = \emptyset$, then lin $\mathcal{C} = \emptyset$ and Phase I gives us an extreme point.

- $z \in \mathcal{C}$ is an extreme point if $z = \alpha\bar{z} + (1-\alpha)\hat{z}$ for $0 < \alpha < 1$ and $\bar{z}, \hat{z} \in \mathcal{C}$ implies that $z = \bar{z} = \hat{z}$.
- $z \in \mathcal{C}$ is a BFS if $\{A_{\cdot j} : l_j < z_j < u_j\}$ are linearly independent.
- $z \in \mathcal{C}$ is a BFS if and only if it is an extreme point.
- $N_{fr} = \emptyset$ implies z is a BFS, hence an extreme point of $\mathcal{C}$.
- Existence of an extreme point implies that lin $\mathcal{C} = \emptyset$.

Phase I result interpretation (when $\nexists$ extreme points)

If $N_{fr} \neq \emptyset$, then lin $\mathcal{C} \neq \emptyset$ and Phase I gives us a boundary point.

- Define $z = (\bar{z}, \hat{z})$ where $\hat{z} = z_{N_{fr}}$. Fix $\hat{z} = 0$.
- Then we have a solution to the following Phase I.

$$\begin{array}{ll} \underset{z}{\text{minimize}} & 0^T z \\ \text{subject to} & Az = a \\ & l \leq z \leq u \\ & \hat{z} = 0 \end{array}$$

- $\bar{z}$ is a BFS in the reduced space of $\mathcal{C}$ where $\hat{z} = 0$, thus an extreme point in that space.

Initial basis setup for starting Lemke's method

From Phase I, we have a nonsingular **B**

$$\mathbf{B}_{\text{PhaseI}} = \begin{bmatrix} A_{\mathcal{A}B} & 0 \\ A_{\mathcal{I}B} & -\mathbb{I}_{\mathcal{I}} \end{bmatrix}$$

where

$$\mathcal{A} : \text{the set of indices of active constraints}$$

$$\mathcal{I} : \text{the set of indices of inactive constraints}$$

So that $A_{\mathcal{A}B}$ is nonsingular.

Initial basis setup for starting Lemke's method

We need to solve a system of equations using complementary pivoting.

$$\begin{aligned} (Mz+q) - \mu r &= A^T u + w - v \\ Az - s &= a \\ 0 \le s &\perp u \ge 0 \\ 0 \le w &\perp z - l \ge 0 \\ 0 \le v &\perp u - z \ge 0 \\ r &\in N_{\mathcal{C}}(z_{\text{PhaseI}}) \end{aligned}$$

If $N_{fr} = \emptyset$,

$$\mathbf{B}_{\text{Lemke}} = \begin{bmatrix} M_{BB} & -A_{\mathcal{A}B}^T & 0 & 0 & 0 \\ M_{LB} & -A_{\mathcal{A}L}^T & -I_L & 0 & 0 \\ M_{UB} & -A_{\mathcal{A}U}^T & 0 & I_U & 0 \\ A_{\mathcal{A}B} & 0 & 0 & 0 & 0 \\ A_{\bar{\mathcal{A}}B} & 0 & 0 & 0 & -I_{\bar{\mathcal{A}}} \end{bmatrix}, \ \text{Bvars} = \begin{bmatrix} z_B \\ u_{\mathcal{A}} \\ w_L \\ v_U \\ s_{\bar{\mathcal{A}}} \end{bmatrix}$$

Initial basis setup for starting Lemke's method

If $N_{\text{fr}} \neq \emptyset$,

$$\mathbf{B}_{\text{Lemke}} = \begin{bmatrix} M_{BB} & M_{BF} & -A_{\mathcal{A}B}^T & 0 & 0 & 0 \\ M_{LB} & M_{LF} & -A_{\mathcal{A}L}^T & -I_L & 0 & 0 \\ M_{UB} & M_{UF} & -A_{\mathcal{A}U}^T & 0 & I_U & 0 \\ A_{\mathcal{A}B} & A_{\mathcal{A}F} & 0 & 0 & 0 & 0 \\ A_{\bar{\mathcal{A}}B} & A_{\bar{\mathcal{A}}F} & 0 & 0 & 0 & -I_{\bar{\mathcal{A}}} \end{bmatrix}, \ \text{Bvars} = \begin{bmatrix} z_B \\ z_F \\ u_{\mathcal{A}} \\ w_L \\ v_U \\ s_{\bar{\mathcal{A}}} \end{bmatrix}$$

If M is invertible in the lineality space of $\mathcal{C}$, then the above matrix is invertible.

Initial pivoting

Solve

$$\begin{bmatrix} M_{BB} & -A_{\mathcal{A}B}^T & 0 & 0 & 0 \\ M_{LB} & -A_{\mathcal{A}L}^T & -I_L & 0 & 0 \\ M_{UB} & -A_{\mathcal{A}U}^T & 0 & I_U & 0 \\ A_{\mathcal{A}B} & 0 & 0 & 0 & 0 \\ A_{\bar{\mathcal{A}}B} & 0 & 0 & 0 & -I_{\bar{\mathcal{A}}} \end{bmatrix} \begin{bmatrix} z_B \\ u_{\mathcal{A}} \\ w_L \\ v_U \\ s_{\bar{\mathcal{A}}} \end{bmatrix} = \begin{bmatrix} -q_B - M_{BL}z_L - M_{BU}z_U \\ -q_L - M_{LL}z_L - M_{LU}z_U \\ -q_U - M_{UL}z_L - M_{UU}z_U \\ b_{\mathcal{A}} - A_{\mathcal{A}L}z_L - A_{\mathcal{A}U}z_U \\ b_{\bar{\mathcal{A}}} - A_{\bar{\mathcal{A}}L}z_L - A_{\bar{\mathcal{A}}U}z_U \end{bmatrix}$$

- Note that z_B and $s_{\bar{\mathcal{A}}}$ are feasible due to Phase I.
- If any of $u_{\mathcal{A}}, w_L$, or v_U is infeasible, then make r basic by increasing μ so that all of them become feasible.

$$r = \left(\sum_{i\in\mathcal{A}} \begin{bmatrix} -A_{iB}^T \\ -A_{iL}^T \\ -A_{iU}^T \end{bmatrix} + \sum_{i\in L} \begin{bmatrix} 0 \\ -I_i \\ 0 \end{bmatrix} + \sum_{i\in U} \begin{bmatrix} 0 \\ 0 \\ I_i \end{bmatrix} \right) \in N_{\mathcal{C}}(z_{\text{Phase I}})$$

Experimental results (LPs)

Some promising results:

Data set	# iterations (Lemke)		Total elapsed time (secs)	
	PathAVI	PATH	PathAVI	PATH
25fv47	3938	3202	0.608037	1.788112
bnl1	592	3230	0.084005	0.616039
pilotnov	3046	$> 10{,}000$	0.668043	> 7.456466
scfxm3	988	4129	0.140008	1.064067
wood1p	336	1325	0.216013	7.120446
woodw	1292	9878	0.652040	27.145696

Table : Solving LP (linear programming) problems using PathAVI and PATH (netlib data sets)

Experimental results (symmetric psd QPs)

Data set	# iterations (Lemke)		Total elapsed time (secs)	
	PathAVI	PATH	PathAVI	PATH
cvxqp1_M	340	1063	0.076004	0.532033
dualc8	4	39	0.008000	0.008001
qscagr25	240	868	0.020001	0.052004
qscfxm3	1072	2021	0.160009	0.504031
qship12l	1399	3246	0.524033	1.188074
cont-101	99	750	18.049127	118.071378

Table : Solving QP (quadratic programming) problems using PathAVI and PATH, Q is symmetric and PSD

- QP problems were taken from "I. Maros, Cs. Meszaros: A Repository of Convex Quadratic Programming Problems, Optimization Methods and Software, 1999"

Experimental results (unsymmetric pd M)

Data set	# iterations (Lemke)		Total elapsed time (secs)	
	PathAVI	PATH	PathAVI	PATH
bnl1	657	$>$ 10,000	0.136008	$>$ 26.065629
capri	296	571	0.016001	0.100006
fit1d	1346	1839	0.156010	0.232014
scsd8	1414	2155	0.936058	3.152197
scfxm3	823	2262	0.212014	5.736358
wood1p	413	915	0.288018	1.440090

Table : Solving AVI problems using PathAVI and PATH, M is unsymmetric PD

- M was randomly generated using MATLAB.

Conclusions

- Treat feasible set $\mathcal{C}$ and $\mathcal{N}_{\mathcal{C}}$ explicitly leads to stronger theory
- Ensure feasibility $\mathcal{C} \neq \emptyset$, and F only evaluated over $\mathcal{C}$
- Works when ∇F is not symmetric
- Can implement theory in large scale setting and get robustness (avoid rank deficiency in initial basis, high accuracy)
- Faster
- Available (subroutine or within GAMS/EMP) - requires CPLEX
- Embed AVI solver in a Newton Method for VI
  - Preprocessing incorporated
  - Each Newton step solves an AVI
  - Hot start critical
  - Nonmonotone pathsearch, watchdogging (another talk)

Splitting Methods

- Suppose $\mathcal{T}$ is a maximal monotone operator

$$0 \in F(z) + \mathcal{T}(z) \qquad (GE)$$

- Can devise Newton methods (e.g. SQP) that treat F via calculus and $\mathcal{T}$ via convex analysis
- Alternatively, can split $F(z) = A(z) + B(z)$ (and possibly $\mathcal{T}$ also) so we solve solve (GE) by solving a sequence of problems involving just

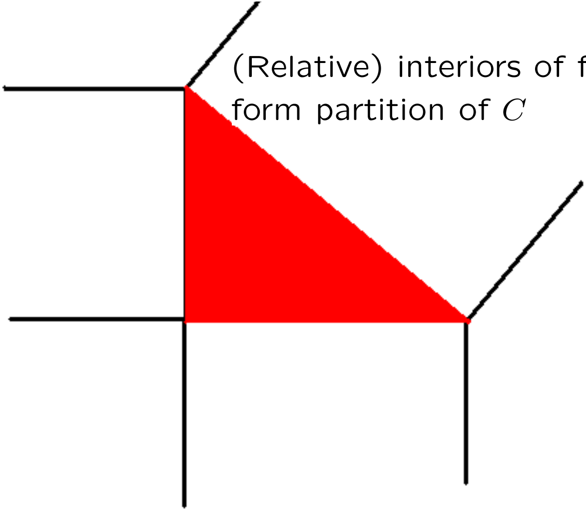
$$\mathcal{T}_1(z) = A(z) \text{ and } \mathcal{T}_2(z) = B(z) + \mathcal{T}(z)$$

where each of these is "simpler"

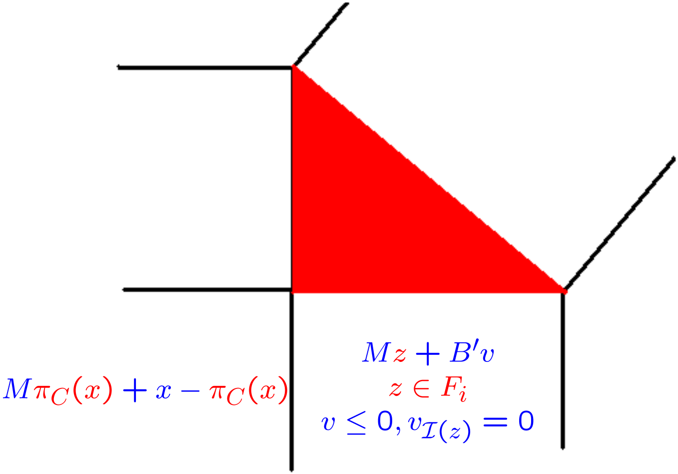
- Forward-Backward splitting:

$$z^{k+1} = (I + c_k T_2)^{-1} (I - c_k T_1) \left(z^k\right),$$

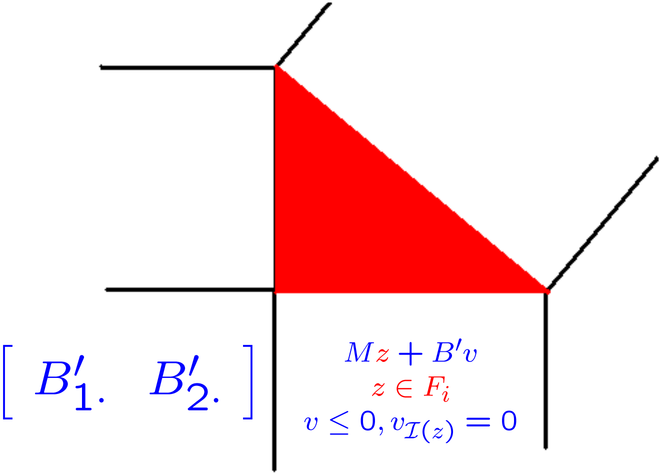
Normal manifold $= \{F_i + N_{F_i}\}$



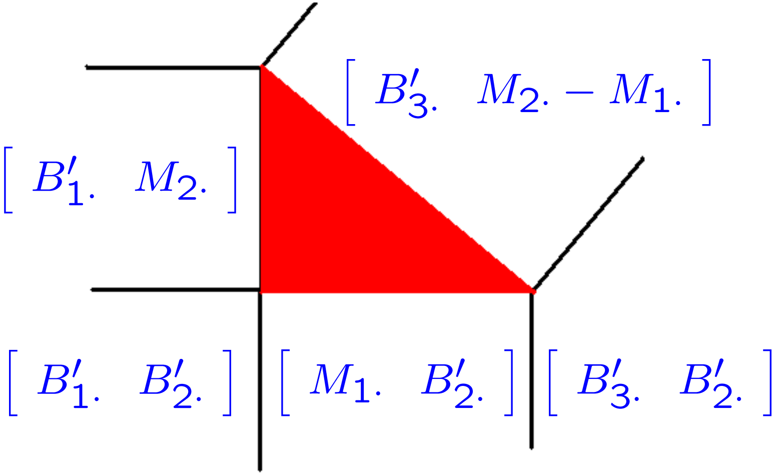
$C = \{z | Bz \geq b\}, N_C(z) = \{B'v | v \leq 0, v_{\mathcal{I}(z)} = 0\}$



$$C = \{z | Bz \geq b\}, N_C(z) = \{B'v | v \leq 0, v_{\mathcal{I}(z)} = 0\}$$

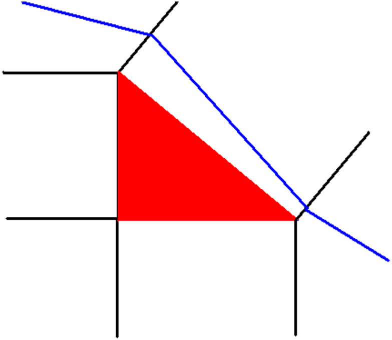


$C = \{z | Bz \geq b\}, F(z) = Mz + q$



Cao/Ferris Path (Eaves)

- Start in cell that has interior (face is an extreme point)
- Move towards a zero of affine map in cell
- Update direction when hit boundary (pivot)
- Solves or determines infeasible if M is copositive-plus on rec(C)
- Solves 2-person bimatrix games, 3-person games too, but these are nonlinear



But algorithm has exponential complexity (von Stengel et al)